

**SYSTEM FOR TRANSACTION PROCESSING
WITH PARALLEL EXECUTION**

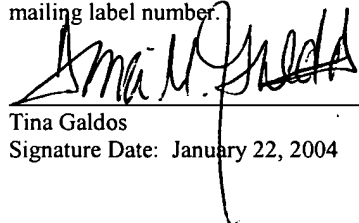
INVENTORS

Alexander J. Somogyi
Adam Messinger
Anno R. Langen

**CERTIFICATE OF MAILING BY "EXPRESS MAIL"
UNDER 37 C.F.R. §1.10**

"Express Mail" mailing label number: EV 385 257 507 US
Date of Mailing: January 22, 2004

I hereby certify that this correspondence is being deposited with the United States Postal Service, utilizing the "Express Mail Post Office to Addressee" service addressed to Mail Stop PATENT APPLICATION, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 and mailed on the above Date of Mailing with the above "Express Mail" mailing label number.



(Signature)

Tina Galdos

Signature Date: January 22, 2004

SYSTEM FOR TRANSACTION PROCESSING WITH PARALLEL EXECUTION

INVENTORS

Alexander J. Somogyi
Adam Messinger
Anno R. Langen

Claim to Priority

[0001] The present application claims the benefit of priority under 35 U.S.C. §119(e) to U.S. Provisional Patent Application entitled "PARALLEL TRANSACTION EXECUTION USING JTA INTERFACE" Serial No. 60/442,319, filed on January 24, 2003, which application is incorporated herein by reference.

Copyright Notice

[0002] A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

Field of the Invention

[0003] The current invention relates generally to XA interactions in Java Transaction API, and more particularly to parallel transaction processing in XA interactions in Java Transaction API.

Background of the Invention

[0004] Java standards for web services are constantly being developed. Concurrently, businesses are building important applications on top of web services infrastructures, such as that available in WebLogic Server from BEA Systems of San Jose, CA. As these applications evolve, they become more complex with more operations to perform. In many implementations, a server may perform transactions involving multiple resources. Transactional guarantees when making updates to the resources are important and should be performed as quickly as possible.

[0005] The WebLogic Server (WLS) Transaction Manager (TM) provides a full-featured two-phase commit transaction engine that implements the Java Transaction API (JTA) specification. Java 2 Enterprise Edition (J2EE) applications deployed on WLS make use of the TM and the JTA interfaces to effect atomic, persistent changes to data that are managed by one or more transactional resources. Transactional resources, such as RDBMS and messaging systems, support the XAResource interface as defined by the JTA specification. A representative server-resource system 100 is illustrated in FIG. 1. In FIG. 1, an environment for a server computer 110 includes server 120, a first resource 130, and second resource 140. Server 120 includes transaction manager 150. Resources are accessed by an application to perform updates to persistent data. The transaction manager is used to ensure that the updates happen atomically with transactional guaranties. This process is described in more detail in “Open Group Distributed Transactions Processing” by Java Inc, incorporated by reference herein.

[0006] WLS internally manages several thread pools that are used to process application logic, administrative requests, inter-server communication and other operations specific to server management. The TM utilizes server threads to perform various transaction-related operations on behalf of the application such as coordinating transactions across multiple servers, transaction timeout processing, transaction recovery processing, and other operations. System 200 of FIG. 2 illustrates a server thread of the prior art. As shown, a server 120 includes a thread 210.

Previously, accessing transactional resources from within a WLS global transaction would cause the TM to perform all resource interactions from within the thread assigned to the application component. Thus, as shown in FIG. 2, the thread 210 includes all interactions such as access resources R1 and R2, two-phase commit processing which includes preparing resources R1 and R2 and committing resources R1 and R2. The two transactional resources are registered and accessible to applications.

[0007] There are several XA interactions that take place between the TM and a resource during the course of a transaction as prescribed by the JTA specification. With reference to FIGs. 1 and 2, an application may begin a transaction and update resources 130 and 140 on server 120. The application logic is dispatched to a server thread 210 running in server 120. The application thread then starts a transaction, updates resources, and commits the transaction. Within the thread 210, the application updates data in first resource 130 and then a second resource 140.

[0008] A method 300 of the prior art for performing a transaction according to the JTA specification is illustrated in FIG. 3. Method 300 begins with start step 305. Next, resources are enlisted at step 310. Resource enlistment occurs when a resource is accessed within the scope of a transaction and it associates the application's changes with a transaction that is under the control of the TM. Thus, when the application commits the transaction, control of the

application thread is given to the TM 150. Next, the participating resources are signaled to prepare to commit in step 320. The TM 150 then sequentially prepares each resource. Each resource is asked by the TM whether the application changes made under the transaction can be made permanent. The TM then determines if all resources are able to commit at step 330. Once all resources have responded to the prepare directive, the transaction manager will either commit or abort the transaction and inform the resource participants of the outcome. If the resources do not all respond positively to the prepare request, then operation continues to step 350 where the TM aborts the transaction. The application may alternatively issue an abort request for the transaction and the TM will only invoke the rollback operation on the resources. If all resources commit at step 330, the TM issues a commit operation signal and the transaction proceeds at step 340. After the TM signals for either a commit or an abort, operation ends at step 365. The prepare, rollback and commit operations are issued in sequence to each locally-accessible resource that is enlisted in the transaction on any given server that is participating in the transaction. If there are additional servers and resources participating in the transaction then each server would invoke the XA operations serially for those participating resources that are locally accessible.

[0009] Performing resource prepare, commit and rollback operations serially as currently done in the prior art can be inefficient since these XA operations may require long processing times. What is needed is a system and method for processing transactions that overcomes the limitations and disadvantages of the prior art.

Summary of the Invention

[0010] In one embodiment of the present invention, a system for utilizing available server threads to process resources and reduce the overall time of performing XA interactions in two-phase commit protocol implemented by the transaction manager (TM) of the present invention. A TM processing XA interactions dispatches interaction commands for multiple resources to idle server threads. In one embodiment, the TM attempts to dispatch all but one of the interaction commands to separate threads. The primary thread then processes the remaining resource command. Any commands relating to dispatch requests that were unable to be dispatched to separate threads due to unavailability are also processed by the primary thread. Once the primary server has processed its interaction commands and received a signal indicating the threads receiving dispatch requests have completed their respective processing of dispatched commands, the next group of commands is processed in a similar manner.

[0011] By utilizing available server threads, it is possible to reduce the overall time spent in XA operations when multiple resources are involved in a transaction by performing the XA operations in parallel. Because the methods of the XAResource interface are synchronous, the use of additional threads is required to obtain concurrency. If there are no idle threads available, these operations will be performed in the application's thread.

Brief Description of the Drawings

[0012] FIGURE 1 is an illustration of a system for performing transactions utilizing servers and their resources in accordance with one embodiment of the present invention.

[0013] FIGURE 2 is an illustration of a serial transaction system having a single thread in accordance with the prior art.

[0014] FIGURE 3 is an illustration of a method for performing a serial transaction in accordance with the prior art.

[0015] FIGURE 4 is an illustration of a parallel transaction system utilizing multiple threads in accordance with one embodiment of the present invention.

[0016] FIGURE 5 is an illustration of a method for performing a parallel transaction in accordance with one embodiment of the present invention.

[0017] FIGURE 6 is an illustration of the average completion time of a transaction in accordance with one embodiment of the present invention.

[0018] FIGURE 7 is an illustration of the average response time of a transaction in accordance with one embodiment of the present invention.

[0019] FIGURE 8 is an illustration of the transaction throughput time in accordance with one embodiment of the present invention.

Detailed Description

[0020] In one embodiment of the present invention, a system of utilizing available server threads to process resources and reduce the overall time for performing XA interactions in two-phase commit protocol implemented by JTA. A TM that processes XA interactions dispatches interaction commands for multiple resources to idle server threads. In one embodiment, the TM attempts to dispatch all but one of the interaction commands to separate threads. The primary thread then processes the remaining resource command. The primary thread may also process any commands relating to dispatch requests that were unable to be dispatched to separate threads. Once the primary server has processed its interaction commands and received a signal indicating the threads receiving dispatch requests have completed their respective processing of dispatched commands, the next group of commands is processed in a similar manner.

[0021] By utilizing available server threads, it is possible to reduce the overall time spent in XA operations when multiple resources are involved in a transaction by performing the XA operations in parallel. Because the methods of the XAResource interface are synchronous, the use of additional threads is required to obtain concurrency. In one embodiment of the present invention, if there are N resources enlisted in a transaction, the TM commit/rollback logic will schedule requests to prepare/commit/rollback for the resources numbered 2 through N, where $N > 1$. This provides the potential of all but one of the resource transactions to be dispatched to other threads. If there are no idle threads available, these operations will be performed sequentially in the application's thread.

[0022] System 400 of FIG. 4 illustrates a server wherein transactions are processed in parallel. System 400 includes server 410, a primary thread 420, T_1 , thread T_X 430, and thread T_Y 440. Primary thread 420 includes resource access commands, a prepare and a commit command for

resource one. Thread T_X 430 includes a prepare command for resource two. Thread T_Y 440 includes a commit command for resource two. As indicated by the time line adjacent to primary thread 420, the commands in the primary thread are processed sequentially in time, but in parallel with commands in thread 430 and thread 440.

[0023] In the present invention, the application thread drives the two phase-commit protocol. If there is more than one local resource participant for the server, the TM will attempt to dispatch resources 2-N to available threads for execution. The request may be executed in the primary thread if there are no idle threads available. After requests for resources 2 through N have been dispatched, the primary thread will perform the XA operation for resource 1 (R_1). After the R_1 operation completes, the primary thread will wait until each dispatched resource request has completed and signaled the primary thread. Thus, as shown in system 400, the prepare R_1 command and prepare R_2 command are performed by threads T_1 and T_X threads, respectively. The TM does not continue the operation of the thread T_1 and process the commit R_1 command until the Prepare R_2 command has completed and signaled the primary thread T_1 .

[0024] The parallel processing of the present invention is further illustrated in method 500 of FIG. 5. Method 500 begins with start step 505. The method of FIG. 5 may be used to invoke several types of XA operations, including resource prepare, commit, and rollback operations. As discussed above, the XA operation type is processed for each resource before processing of the next XA operation begins. The XA operation to be performed depends upon the phase in the two-phase commit protocol being processed. Step 510 signifies that an XA operation associated with a phase of the two-phase commit protocol is selected to be processed. The number and identity of available threads is determined at step 520. In one embodiment, the transaction manager, or TM, provides tasks to the thread pool manager. The thread pools maintained by the

server system are then checked to determine which threads are idle. Information for idle threads within the thread pool are determined by the thread manager. Next, the thread pool manager determines if there is an available thread for the current resource XA transaction at step 530. If an idle thread does not exist, operation continues to step 535 wherein the request may be immediately invoked in the current thread or caller's thread as a blocking operation. In this case, the thread pool manager processes the work in the current thread. Thus, when there are no idle threads, the transaction operations are performed in sequence. In yet another embodiment, a dedicated thread pool is used for parallel transaction operations. In this case, instead of using a default thread pool to dispatch transaction operations for idle threads, operations are dispatched to a separate, dedicated pool used for transaction operation. After the operation is invoked in the current thread, operation of method 500 continues to step 550.

[0025] If a thread is available at step 530, operation continues to step 540 wherein the thread manager dispatches the XA operation to the available thread. Operation then continues to step 550. In one embodiment, only the XA operation for each resource is dispatched. For example, as shown in system 400 of FIG. 4, prepare to commit transactions are handled at the time indicated by indicator 450. At that time, the prepare operations are handled by primary thread 420 for the first resource and thread 430 for the second resource.

[0026] At step 550, if the TM has attempted to dispatch the current XA operation of all N-1 resources in the transaction, operation continues to step 560. If there are more XA operations, the next resource is addressed as indicated in step 555 and operation continues for the resource at step 520. At step 560, the primary thread then processes the XA operation for resource 1.

[0027] After the primary thread has processed the XA operations in step 560, the primary thread determines if the threads that received dispatches in step 560 have signaled the primary thread at

step 570. In one embodiment, the primary thread determines if the respective threads have signaled the primary thread to indicate their respective commands have been processed. In one embodiment, operation remains at step 570 until all threads receiving dispatched commands have signaled the primary thread. In another embodiment, the primary thread will wait for a pre-determined period of time during which the threads receiving dispatches may signal the primary thread. In this embodiment, if the threads do not signal the primary thread within the time period, a time-out error is logged. If the TM is processing the prepare phase of the two phase commit protocol, then the transaction may be aborted. Once all threads processing dispatched requests have signaled the primary thread, the TM reports results of the XA operation at step 580. Operation then ends at step 595.

[0028] In one embodiment of the present invention, the signaling mechanism is based on a Java class that maintains a counter of the number of signal events anticipated. The primary thread will invoke a wait method that blocks the caller until the object is signaled the set number of times. The modified prepare/commit/rollback algorithms sets the expected number of signal events to be N-1. Each request that is dispatched to a thread for execution is given a reference to the synchronization object. Upon completing the XA operation the thread will invoke the object's signal method before terminating.

[0029] FIGs. 6-8 illustrate data corresponding to traits of the present invention. In FIGs. 6-8, data plots corresponding to three types of XA interaction processing implementations are provided. For each plot, the server is assumed to have three resources, a one second prepare to commit time, and a one second commit time. The darkest plot containing interspersed diamond icons represents the results of 11 free default queue threads and 11 dedicated XA threads. The second darkest plot containing interspersed squares represents 22 free default queue threads.

This second darkest plot most resembles the serial processing technique of the prior art. The lightest plot containing interspersed triangles includes 22 free default queue threads, wherein parallel transactions are implemented in the default queue threads if one or more threads are idle.

[0030] FIG. 6 illustrates a plot of the average transmission completion time in seconds for varying numbers of client threads used by applications. For low numbers of client threads being used, the parallel transaction processing system provides for much faster processing and better thread utilization than the serial processing technique. As the number of threads in use by a server increase, the average transaction completion time improvement of the parallel processing method over the serial processing method slowly decreases. FIG. 7 illustrates a plot of the average response time. For low numbers of client threads being used, the parallel transaction processing system provides for much faster average response time than the serial processing technique. As the number of threads in use by a server increase, the average response time improvement of the parallel processing method over the serial processing method slowly decreases. FIG. 8 illustrates a plot of the transaction throughput characteristic of the three processing systems. For low numbers of client threads being used, the parallel transaction processing system provides for transactions completed per second than the serial processing technique. As the number of threads in use by a server increase, the transactions per second improvement of the parallel processing method over the serial processing method slowly decreases.

[0031] As indicated in FIGs. 6-8, server performance improves with the use of the parallel XA feature when client loads are light to moderate. This is due to the utilization of otherwise idle threads. For heavy client loads, where all server threads are being utilized for processing client

requests, the behavior and performance is similar to that of serialized XA in that the primary thread processes most of the XA related interactions.

[0032] In one embodiment of the present invention, available server threads are utilized to process resources and reduce the overall time of performing XA interactions in two- phase commit protocol implemented by the TM of the present invention. A TM that processes XA interactions dispatches interaction commands for multiple resources to idle server threads. In one embodiment, the TM attempts to dispatch all but one of the interaction commands for current command. The primary thread then processes the remaining resource command as well as any commands relating to dispatch requests that failed. Once the primary server has processed its interaction commands and received a signal indicating the threads receiving dispatch requests have completed their respective processing of dispatched commands, the next group of commands is processed in a similar manner.

[0033] Other features, aspects and objects of the invention can be obtained from a review of the figures and the claims. It is to be understood that other embodiments of the invention can be developed and fall within the spirit and scope of the invention and claims.

[0034] The foregoing description of preferred embodiments of the present invention has been provided for the purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise forms disclosed. Obviously, many modifications and variations will be apparent to the practitioner skilled in the art. The embodiments were chosen and described in order to best explain the principles of the invention and its practical application, thereby enabling others skilled in the art to understand the invention for various embodiments and with various modifications that are suited to the particular use contemplated. It is intended that the scope of the invention be defined by the following claims and their equivalence.

[0035] In addition to an embodiment consisting of specifically designed integrated circuits or other electronics, the present invention may be conveniently implemented using a conventional general purpose or a specialized digital computer or microprocessor programmed according to the teachings of the present disclosure, as will be apparent to those skilled in the computer art.

[0036] Appropriate software coding can readily be prepared by skilled programmers based on the teachings of the present disclosure, as will be apparent to those skilled in the software art. The invention may also be implemented by the preparation of application specific integrated circuits or by interconnecting an appropriate network of conventional component circuits, as will be readily apparent to those skilled in the art.

[0037] The present invention includes a computer program product which is a storage medium (media) having instructions stored thereon/in which can be used to program a computer to perform any of the processes of the present invention. The storage medium can include, but is not limited to, any type of disk including floppy disks, optical discs, DVD, CD-ROMs, microdrive, and magneto-optical disks, ROMs, RAMs, EPROMs, EEPROMs, DRAMs, VRAMs, flash memory devices, magnetic or optical cards, nanosystems (including molecular memory ICs), or any type of media or device suitable for storing instructions and/or data.

[0038] Stored on any one of the computer readable medium (media), the present invention includes software for controlling both the hardware of the general purpose/specialized computer or microprocessor, and for enabling the computer or microprocessor to interact with a human user or other mechanism utilizing the results of the present invention. Such software may include, but is not limited to, device drivers, operating systems, and user applications.

[0039] Included in the programming (software) of the general/specialized computer or microprocessor are software modules for implementing the teachings of the present invention, including, but not limited to, parallel processing of XA interactions in a JTA environment.